REMARKS

I. Status of the Claims/Amendments

Claims 1-24, 27-35, 43, 44, 48-54 and 58-68 are pending. Claims 1, 22, 23, 27 and 68 are currently amended. Claims 69 to 75 are new claims.

II. Objection to Drawings

The replacement Fig. 2 has been amended to improve legend legibility.

III. Compliance with Sequence Listing Requirement

A substitute Sequence Listing is submitted along with a computer disk containing the same Sequence Listing in computer readable form.

IV. Provisional Double Patenting Rejection

Claims 1-24, 27-35, 43, 44, 58-54 and 58-68 are provisionally rejected as claiming the same invention as that of claims 1-11,26, 27, 36-43, 44, 45 and 69-74 of co-pending Application No. 09/917,384. Applicants have reviewed both sets of sequence listings and find them to be completely different. The sequences used for the alignments by the Examiner are all sequences from the present application. Applicants note that the present application and the co-pending application No. 09/917,384 have almost identical inventors, and the application serial numbers differ only in the last digit. It is possible that the issue raised by the examiner has resulted from a clerical error. Applicants are resubmitting a substitute sequence listing for each of the applications to cure the problem. Since the claims are directed to different sequences, Applicants respectfully request that the Office withdraw the provisional double patenting rejection of claims 1-24, 27-35, 43, 44, 58-54 and 58-68.

V. Rejection Under 35 U.S.C. § 101

Claims 22-24 and 68 stand rejected under 35 U.S.C. §101 for lack of utility. We respectfully traverse.

The present application discloses the generation of various GuxA derivatives that contain GuxA and heterologous polypeptides. The specification assets utility for these compositions, e.g., "[p]referred heterologous polypeptides include those that facilitate purification, oligomerization, stability, or the secretion of the GuxA polypeptide." (p. 19, lines 14-16). Also discussed in the specification, "GuxA polypeptides can be fused to heterologous polypeptides to facilitate purification. Many available heterologous peptides (peptide tags) allow selective binding of the fusion protein to a binding partner. Non-limiting examples of peptide tags include 6-His, thioredoxin, hemaglutinin, GST, and the OmpA signal sequence tag. A binding partner that recognizes and binds to the heterologous peptide can be any molecule or compound, including metal ions (for example, metal affinity columns), antibodies, antibody fragments, or any protein or peptide that preferentially binds the heterologous peptide to permit purification of the fusion protein." (p. 20, lines 16-22).

In addition, "GuxA polypeptides can be modified to facilitate formation of GuxA oligomers. For example, GuxA polypeptides can be fused to peptide moieties that promote oligomerization, such as leucine zippers and certain antibody fragment polypeptides, for example, Fc polypeptides. General techniques for preparing fusion proteins are known, and are described, for example, in WO 99/31241 and in Cosman et al., 2001 *Immunity* 14:123-133. Fusion to an Fc polypeptide offers the additional advantage of facilitating purification by affinity chromatography over Protein A or Protein G columns. Fusion to a leucine-zipper (LZ), for example, a repetitive heptad repeat, often with four or five leucine residues interspersed with other amino acids, is described in Landschultz et al., 1988 *Science*, 240:1759." (p. 20, lines 24-32).

Therefore, Claims 22-24 and 68, as amended, have numerous utilities that are well supported by the instant specification and by common knowledge in the art. Applicants respectfully request that the Office withdraw the rejection of claims 22-24 and 68.

The Examiner asserts, particularly, that the specification does not teach that the claimed heterologous combination will have any particular activity or utility. That finding is directly controverted by the preceding discussion where, for example, the fusion peptides may be used to facilitate protein purification. There is also the teaching of Example 4 which discloses reshuffling of various domains of different hydrolases to address specific needs, such as to facilitate purification, to enhance substrate binding capability or to increase catalytic activity of the enzymes.

The Office seems to state a case for lack of utility by inoperability. This is not the case where, for example, the art recognizes that fusion peptides may be constructed to facilitate purification and to provide multiple domain peptides. "Where the mode of operation alleged can be readily understood and conforms to the known laws of physics and chemistry, operativeness is not questioned, and no further evidence is required." *In re Chilowsky*, 229 F.2d 457, 462, 108 USPQ 321, 325 (CCPA 1956). Therefore, we respectfully request the Office to withdraw the lack of utility rejection.

VI. Rejection of Claims 22-24 and 68 Under 35 U.S.C. § 112, First Paragraph

Claims 22-24 and 68 also stand rejected under 35 U.S.C. §112, first paragraph. The Examiner stated that "[s]ince the claimed invention is not supported by either a asserted utility or a well established utility for the reasons set forth [in the rejection under 35 U.S.C. §101], one skilled in the art clearly would not know how to use the claimed invention."

As discussed above in Section V, numerous utilities are asserted in and supported by the specification. Therefore, one of ordinary skill in the art would know how to use the cellulase claimed herein in at least those capacities. For the foregoing reasons,

Applicants respectfully requests withdrawal of the rejection of claims 22-24 and 68 under 35 U.S.C. § 112, first paragraph.

VII. Rejection of Claims 27, 28, 35, 43, 44, 48-54 and 63-67 under 35 U.S.C. § 112, First Paragraph

Claims 27, 28, 35, 43, 44, 48-54 and 63-67 stand rejected under 35 U.S.C. §112, first paragraph as failing to comply with the written description requirement. The Examiner asserted that the "claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention."

Applicants respectfully traverse because the functionalities of GH6 and GH12 hydrolases are defined in the specification. For instance, the instant specification states that, "[t]he GH6 domain family includes a number of cellbiohydrolases, for example, exocellbiohydrolase A isolated from *Cellulomonas fimi*, and exoglucanase E3 isolated from *Thermobifida fusca*. The GH6 members degrade substrate using an inverting mechanism. The G12 domain family includes a number of endoglucanases, for example, endo-1,4-glucanase isolated from *Streptomyces lividans*, and endo-1,4-glucanase S cellulase 12A isolated from *Streptomyces sp.* 11AG8. The GH12 members degrade substrate using a retaining mechanism." (p. 17, lines 10-15). Even though the functionality is described, claim 27 has been amended to recite --domain--, as opposed to the former "functionality" to clarify that the domain is being claimed. The functionality is inherent to the domain.

The Examiner further pointed out that the terms "glycosyl hydrolase family 6 and glycosyl hydrolase family 12" as used in claim 27 are not defined. However, the terms "GH6" and "GH12" are the acronyms for "glycosyl hydrolase family 6" and "glycosyl hydrolase family 12," respectively. This notion is further supported by the instant specification which states on page 15 at lines 14-21:

> Glycoside hydrolase enzymes are classified into glycoside hydrolase (GH) families based on significant amino acid similarities within their catalytic domains. Enzymes having related catalytic domains are grouped together within a family, (Henrissat et al. (1991) supra, and Henrissat et al. (1996), Biochem. J. 316:695-696), where the underlying classification provides a direct relationship between the GH domain amino acid sequence and how a GH domain will fold. This information ultimately provides a common mechanism for how the enzyme will hydrolyse the glycosidic bond within a substrate, *i.e.*, either by a retaining mechanism or inverting mechanism (Henrissat., B, (1991) supra)."

As shown by the above passage, the designations GH6 and GH12 are well known and readily understood. "A specification need not disclose what is well known in the art."*Genentech, Inc. v. Novo Nordisk A/S*, 108 F.3d 1361, 1366, 42 USPQ2d 1001, (Fed. Cir. 1997). The Declaration of William H. Adney attests to the well-know nature of GH6 and GH12 domains. The cited Henrissat articles are well known and frequently used as the basis for classification in this field. The notion that "glycosyl hydrolase family 6 and glycosyl hydrolase family 12" and "GH6 and GH12" refer to the same thing is also supported by prior art cited in the instant application. In grouping different glycoside hydrolase families into clans, one of the cited references, Henrissat et al., used "GH" families to represent "glycoside hydrolase" families (See Table 2 of Henrissat et al., *Biochem. J.*, 1996, 316: 695-696). Thus, "glycosyl hydrolase family 6 and glycosyl hydrolase family 12" and "GH6 and GH12" are used interchangeably in the art. Furthermore, both GH6 and GH12 families, as well as representative enzymes of each family, are disclosed in the specification. Applicants respectfully request withdrawal of the rejections of these claims under 35 U.S.C. §112, first paragraph.

VIII. Rejection Under 35 U.S.C. § 102(b)

Claims 27, 35 and 43 stand rejected under 35 U.S.C. § 102(b) as being anticipated by Sakon *et al.* The Examiner maintained that because applicants' proposed amendment to claim 27 was rejected for introducing new matter, claims 27, 35 and 43 remain anticipated by Sakon *et al.* Although Sakon teaches an amino acid sequence that is 72.1% identical with SEQ ID NO: 8, it describes a GH5 family enzyme (PAGE 10649,

first column, first full paragraph). Claim 27 has been amended to recite the functionality of at least one of family GH6 and GH12, and this distinguishes Sakon. The recitation of GH6 and/or GH12 domains in claim 27 distinguishes Sakon. Claims 35 and 43 depend from claim 27 and are also distinguishable from Sakon.

IX. Rejection Under 35 U.S.C. § 102(b)/103(a)

Claims 1-21, 27-35, 43, 44, 48, 54 and 58-67 stand rejected under 35 U.S.C. §102(b) as anticipated by or, in the alternative, under 35 U.S.C. § 103(a) as obvious over either of Tucker *et al* (AA)., Adney *et al.* (AB), Himmel *et al.* (AC), Lastick *et al.* (AD) or Baker *et al.* (AJ). Applicants traverse because the cited references all disclose enzymes that are distinct from the presently claimed GuxA.

The rejection is speculative where the calculated molecular weight of GuxA is 124,597 daltons, and no proteins of this weight are reported in the cited references (see Declaration of William H. Adney). The cited references AB and AC disclose Endo I and Endo II, respectively, weighing from 57,429 daltons to 74,580 Daltons and between about 50,000 to 70,000 daltons. The Office speculates in this case that multimers may exhibit higher molecular weights in size-exclusion columns than on SDS PAGE.

Applicants agree with the Examiner in that the apparent molecular weight estimated from size-exclusion column could at times be higher than a protein's theoretical molecular weight. However, even taking into consideration possible errors in human experimentation, Applicants are not aware of any reports in the literature, where a protein with a theoretical molecular weight of 124 Kd would be showing an apparent molecular weight of between 50-80 Kd either on SDS PAGE or in a size-exclusion column. Applicants can imagine but one situation where a protein would be exhibiting a molecular weight almost 50 Kd lower than its calculated molecular weight; namely, when the protein is degraded. Indeed, if the enzymes taught in the cited references are merely degraded fragments of the presently claimed GuxA protein, Applicants fail to see why successful isolation of a full-length enzyme amid numerous failed attempts by others to isolate an undegraded protein is not patentable within the meaning of 35 U.S.C. §§ 102

and 103. Partial protein sequence information for these shorter peptides confirms, generally, that they are not degraded GUXA sequences.

Lastick (AD) reports two DNA fragments obtained from *A. cellulolyticus* that are used to express cellulases in *E. coli*. One is a 2kb BamHI fragment and the other is a 3.7kb PvuI fragment. Fig. 1 shows polyacrylamide gel electrophoresis showing activity of the protein bands by the zymogram method (col. 4, lines 2-16). The comparison of Fig. 1 shows differences in electrophoretic migration on different tracks between the "low MW endoglucanase," the "high MW endoglucanase,: and lysate from clone $DH5_{OZ}$(pMTL2-2) incorporating the 2 kb fragment. No sequence is provided; however, positional activity is not analogous to either the low MW or the high MW endoglucanase where the clone $DH5_{OZ}$ has an even lower molecular weight than does the low MW endoglucanase. Thus, GUXA is not the same as the polypeptide encoded by the 2 kb fragment.

It is not even ultimately interesting to observe that GUX A has an approximate length of 3.7 kb, as reported in SEQ ID NO 2 and that this corresponds to the 3.7 kb fragment length reported in Lastick (AD). This is because Lastick (AD) reported an internal BamHI site to that other fragment, whereas SEQ ID NO. 2 does not contain a BamHI site. Thus, the presently claimed GUXA is not the same as the polypeptide encoded by the 3.7 kb fragment reported by Lastick (AD).

Lastly, a high molecular-weight endoglucanase of from 156,600 to 203,400 daltons is taught by Tucker et al. (AA). Tucker (AA) describes a high molecular weight cellulase that exceeds the molecular weight of GuxA. Applicants acknowledge the Office position that more than one cellulases may aggregate which may skew the estimated molecular weight obtained from a size-exclusion column. Applicants also acknowledge the Office position that posttranslational modifications may sometimes distort the apparent molecular weight obtained from SDS-PAGE. However, the apparent difference in molecular weight and the thermal inactivation profile of the Tucker enzyme and GuxA suggest that they are different proteins. In essence, the Office imposes upon Applicants the undue burden of proving a negative and Applicants respectfully maintain that the

burden is on the Office to establish a §102 anticipation or §103 *prima facie* obviousness case where the available evidence does not support the Office position.

Even if the enzymes taught in the cited references are the same as the presently claimed GuxA, the present claims, as amended, are distinguishable from the teachings of the prior art. All cited references taught substantially purified cellulose hydrolases from *Acidothermus cellulolytics*. Nothing is mentioned in the cited references on cloning the genes coding for the hydrolases and express the enzymes in large scale in an organism other than *Acidothermus cellulolyticus*. Amended Claim 1 recites GuxA that is heterologously expressed in an organism other than *Acidothermus cellulolyticus*. Such amendment limits the present claims to recombinant hydrolases and thus distinguishes themselves from the teachings of the prior art which invariably teaches purification of the naturally occurring enzymes from *Acidothermus cellulolyticus*.

As the Declaration by William Adney points out, at the time of the present invention, there was no practicable methodology to sequence the full length of such a long protein as GuxA. In fact, even Edman degradation which was routinely used to sequence the N-terminus of a protein, had an important limitation as to the purity and length of the protein. Even today, it is not practical if not impossible to sequence a long polypeptide such as GuxA. Therefore, at the time of the present invention, one of ordinary skill would not even attempt to directly sequence the Tucker protein preparation in order to obtain the sequence of the endoglucanase because of the low expectation of success. As a matter of fact, molecular cloning of the gene coding for the enzyme was a more practicable approach to obtain the sequence of the enzyme (see Adney Declaration).

As the Federal Circuit has repeatedly held, knowledge of the sequence of a protein, without more, is not sufficient to anticipate or render cloning and sequencing of the encoding gene obvious. *See, e.g., In re Deuel*, 51 F.3d 1552, 34 U.S.P.Q.2d (BNA) 1210 (Fed. Cir. 1995). In the instant case, discovery of the gene makes heterologous expression possible, and the resulting enzyme is modified by the host that is used to express the enzyme. Therefore, even if what the Examiner assumes is true (and here we

make no such admission) none of the cited references anticipate or render obvious the sequencing of the GuxA gene and the deduced protein sequence.

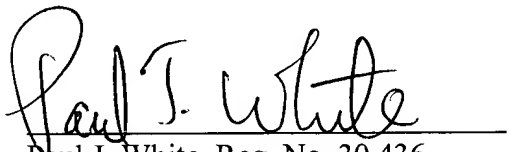
X. Conclusion

Applicants' attorney respectfully solicits a Notice of Allowance in this application. The Commissioner is authorized to charge any additionally required fees to deposit account 14-0460. Should the Examiner have any questions, comments, or suggestions that would expedite the prosecution of the present case to allowance.

Applicants' undersigned representative earnestly requests a telephone call at (303) 384-7575.

Respectfully Submitted,

Date: 8/12/04



Paul J. White, Reg. No. 30,436
Senior Patent Counsel

National Renewable Energy Laboratory
1617 Cole Boulevard
Golden, CO. 80401
(303) 384-7575

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s):	Shi-You Ding et al.		
Serial No.:	09/917,383	Group No.:	1652
Filed:	July 28, 2001	Examiner:	Patterson, Charles L. Jr.
For:	THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS CELLULOLYTICUS	Confirmation No.	9967

CERTIFICATE OF MAILING

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being deposited in the United States Postal Service on the date shown below with sufficient postage for first class mail, in an envelope addressed Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

August 13, 2004
Date

Brenda E. Brantley
Brenda E. Brantley

Mail Stop RCE

Commissioner for Patents

P.O. Box 1450

Alexandria, VA 22313-1450

Sir:

RULE 132 DECLARATION OF WILLIAM S. ADNEY

1. This Declaration is provided to present certain information for the Examiner's review. This information traverses the written description and anticipation and obviousness rejections found in the Office Action dated March 3rd, 2004.

2. My name is William S. Adney and I am named as a coinventor in this application. I am employed by the National Renewable Energy Laboratory (NREL) in

Golden, Colorado. NREL is the assignee of the invention. My job title is that of Senior Scientist II.

3. Exhibit A attached to this Declaration is my brief curriculum vitae summarizing my educational background and work experience.

4. Exhibit B attached to this Declaration is a list of publications and presentations where I am an author, coauthor, or coinventor.

5. I have reviewed the present claims and the Office Action dated March 3rd, 2004.

6. Exhibit C attached to this Declaration is a paper by Henrissat and Bairoch, (*Biochem. J.*, 1996, 316: 695-696). The paper is well known and frequently cited as the basis for classification of glycosyl hydrolases in the field. In grouping different glycoside hydrolase families into clans, Table 2 of the paper used "GH" families to represent "glycoside hydrolase" families. Therefore, the terms "glycosyl hydrolase family 6 and glycosyl hydrolase family 12" and "GH6 and GH12" are used interchangeably in the art.

7. The present application discloses a novel hydrolase with a theoretical molecular weight of 124,597 Daltons. No other hydrolases of similar size have been previously reported either by our group or others in the field. The hydrolases disclosed in references AB and AC have estimated molecular weights of between about 50,000 to 80,000 Daltons. In my experience, a protein with a calculated molecular weight of 124,597 Daltons does not co-migrate on an SDS-PAGE or in a size-exclusion column with a marker protein of less than 80,000 Daltons. In other words, unless the proteins disclosed in references AB or AC were degraded, their biochemical characteristics

support the conclusion that they are distinguishable from the claimed GuxA hydrolase of the present application.

8. The present invention solves the problem of obtaining large amount of cellulose hydrolases for degrading bio-mass. Although several groups, including ourselves, had identified a number of hydrolases from *Acidothermus cellulolyticus* prior to the present invention, the inherently slow grow rate of *A. cellulolyticus* hampers large-scale production of the enzymes (see page 2, lines 18-30 of present application). Before one can express large amount of the hydrolases heterologously, the gene coding for the enzymes have to be cloned. The present application discloses and claims cloning of a hydrolase gene, GuxA, and the heterologous expression of the GuxA protein in an organism other than *A. cellulolyticus*.

9. At the time of the present invention, no methodology was available to obtain the full-length sequence of a protein as large as GuxA. One important tool available at the time to obtain N-terminal sequence of a protein is Edman Degradation. However, factors such as sample purity and instrument sensitivity continue to limit the utility of Edman Degradation. In fact, were it not for the lack of powerful tools for full-length protein sequencing, whole proteom (all the proteins in a cell) sequencing, rather than the now well-known whole genome sequencing, would have taken the center stage of modern biology. It is precisely because of the extreme difficulty of obtaining protein sequence directly from biochemically purified cell lysates that we took on the challenge of molecular cloning in order to heterologously express the protein in a fast-growing organism other than *A. cellulolyticus*.

10. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under 18 U.S.C. § 1001, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: 5 Aug 2004

By: William S Adney

William S. Adney

Biochem. J. (1996) 316, 695–696 (Printed in Great Britain)

BIOCHEMICAL JOURNAL LETTERS

Updating the sequence-based classification of glycosyl hydrolases

A classification of glycosyl hydrolases based on amino-acid-sequence similarities was proposed in this Journal a few years ago [1]. This classification originated from the analysis of ~ 300 sequences and their grouping into 35 families designated 1–35. Because such a classification is necessarily sensitive to the sample, it was anticipated that it was incomplete and that new families would be determined when additional sequences would become available. When the number of glycosyl hydrolase sequences reached ~ 480, ten additional families (designated 36–45) could be defined and were added to the classification [2]. There are at present over 950 sequences of glycosyl hydrolases in the data-banks (EMBL/GenBank and SWISS-PROT). Their analysis shows that the vast majority of the ~ 470 additional sequences that have become available since the last update could be classified in the existing families. However, several sequences not fitting the existing families allow the definition of new families (designated 46–57) (Table 1). When the several present genome sequencing projects have reached completion, the number of

Table 1 New families in the classification of glycosyl hydrolases

Family	Enzyme	Organism	SWISS-PROT	EMBL/GenBank
46	Chitosanase	*Bacillus circulans* MH-K1	P33673	D10624
46	Chitosanase	*Streptomyces* sp. N174	P33665	L07779
46	Chitosanase	*Nocardioides* sp.	P48846	L40408
47	α-Mannosidase	*Drosophila melanogaster*		X82641
47	α-Mannosidase 9	Human	P33908	X74837
47	α-Mannosidase	Mouse	P39098	U03458
47	α-Mannosidase	Mouse	P45700	U04299
47	α-Mannosidase	*Penicillium citrinum*		D45839
47	α-Mannosidase	Rabbit	P45701	U04301
47	α-Mannosidase	*Saccharomyces cerevisiae*	P32906	M63598; Z49631
47	Open reading frame	*Caenorhabditis elegans*		Z47073
48	Cellulase CelS	*Clostridium thermocellum*	P38686	S56455
48	Cellulase CelF	*Clostridium cellulolyticum*	P37698	U30321
48	Open reading frame	*Caldocellum saccharolyticum*	P22534	M36063
48	Cellobiohydrolase B	*Cellulomonas fimi*		L38827
49	Dextranase	*Arthrobacter* sp.	P39652	D00834
49	Dextranase	*Penicillium minioluteum*	P48845	L41562
50	Agarase A	*Vibrio* sp.	P48839	D14721
50	Agarase B	*Vibrio* sp.	P48840	D21202
51	Arabinofuranosidase A	*Aspergillus niger*	P42254	L29005
51	Arabinofuranosidase	*Streptomyces lividans*		U04630
52	β-Xylosidase	*Bacillus stearothermophilus* 236	P45704	U15984
52	β-Xylosidase	*Bacillus stearothermophilus* 21	P45702	D28121
53	Galactanase 1	*Aspergillus aculeatus*	P48842	L34599
53	Open reading frame	*Bacillus polymyxa*	P48843	L03425
53	Galactanase	*Pseudomonas fluorescens*	P48841	X91885
54	Arabinofuranosidase B	*Aspergillus niger*	P42255	X74777
54	Arabinofuranosidase/xylanase	*Trichoderma koningii*	P48792	U38661
55	Exo-1,3-β-glucanase	*Cochliobolus carbonum*	P49426	L48994
55	Endo-1,3-β-glucanase	*Trichoderma harzianum*		X84085
56	Hyaluronidase	*Apis mellifera*	Q08169	L10710
56	Hyaluronidase	*Cavia porcellus*	P23613	X56332
56	Hyaluronidase	*Dolichovespula maculata*	P49 371	L34548
56	Hyaluronidase	Human	P38567	L13781
56	Hyaluronidase	*Macaca fascicularis*	P38568	L13780
56	Hyaluronidase	Mouse	P48794	U33958
56	Hyaluronidase	Rabbit	P38566	U09183
56	Hyaluronidase	*Vespula vulgaris*	P49370	L43562
57	α-Amylase 1	*Dictyoglomus thermophilum*	P09961	X07896
57	α-Amylase	*Pyrococcus furiosus*	P49067	L22346

```
********
Family 6
********

Description: Endoglucanases (EC 3.2.1.4) and
             cellobiohydrolases (EC 3.2.1.91).
PROSITE: PDOC00563
3D structure status: Available
Reaction stereochemical outcome: Inverted anomeric configuration
Catalytic nucleophile/base: Asp (experimental)
Catalytic proton donor: Asp (experimental)
Clan: None
Known taxonomic range: eukaryotae, prokaryotae.
Note: formerly known as cellulase family B.

GUNA_CELFI (P07984), GUNB_FUSOX (P46236), GUNA_MICBI (P26414),
GUNA_STRRA (P33682), GUN1_STRSQ (P13933), GUN2_TERFU (P26222),
GUX2_TRIRE (P07987)
```

Figure 1 Example of a section of the glycosyl hydrolase classification document

The text in **bold** denotes hypertext links to other electronic servers (ENZYME, PROSITE and SWISS-PROT).

glycosyl hydrolase sequences will probably increase dramatically. There are two major problems with keeping the classification: (i) how to make it available *in toto*, and (ii) how to disclose the new families when they are discovered. One way is the publication in scientific journals of papers whose interest decreases as they become progressively similar to stamp collections. Another way is the use of more adapted media such as electronic databases.

We are happy to announce that a permanently updated version of the classification is now available through the ExPASy WWW server [3] at the URL: 'http://expasy.hcuge.ch/cgi-bin/lists?glycosid.txt'. For each family of glycosyl hydrolases, a section of the document exists (Figure 1) that briefly lists the main enzymes that belong to this family. This section includes links to the relevant SWISS-PROT [4] protein-sequence entries. Links are also provided to the relevant EC numbers in the ENZYME [5] nomenclature database as well as to PROSITE [6] entries (which currently exist for more than half of the known glycosyl hydrolase families). This electronic classification should answer the need for rapid updates and availability *in toto* or family by family and allow users to navigate seamlessly between various types of network resources providing information on these enzymes.

There are two major mechanisms for glycosyl hydrolases, leading to overall retention or inversion of the stereochemistry at the cleavage point [7]. The mechanism appears to be conserved within each family [8]. The following families have been found to act with a retaining mechanism : 1, 2, 5, 7, 10, 11, 12, 13, 16, 17, 22, 30, 31, 32, 33, 34, 35, 39 and 42 (the mechanism of families 30, 35 and 42 was inferred from sequence similarities [9]). The inverting mechanism prevails in families 6, 8, 9, 14, 15, 19, 24, 37, 43, 44, 45, 46, 47 and 48. The electronic classification indicates the type of mechanism for each family where it is known.

The three-dimensional structure is now known for at least one member of families 1, 2, 5, 6, 7, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 33, 34, 45 and 46 (for a review, see [10]). The availability of a three-dimensional structure is indicated in the electronic version of the classification.

A 'clan' is a group of families that are thought to have a common ancestry and are recognized by significant similarities in tertiary structure together with conservation of the catalytic residues and catalytic mechanism. The growing number of three-dimensional structures solved for glycosyl hydrolases and/or improved sequence comparison strategies have revealed the relationship between some glycosyl hydrolases families which can be grouped in clans (Table 2).

Table 2 Clan grouping of glycosyl hydrolase families

Clan	Families grouped	Reference
GH-A	1, 2, 5, 10, 17, 30, 35, 39, 42	[9,11]
GH-B	7, 16	[12]
GH-C	11, 12	[13]
GH-D	27, 36	[14,15]
GH-E	33, 34	[16]

The grouping into clans is also indicated in the electronic classification. Families 19, 22, 23 and 24 have been proposed to be related on the basis of weak local three-dimensional similarities [17]. However, since family 22 acts with a retaining mechanism while families 19 and 24 use the inverting mechanism, and since the mechanism of family 23 is not yet known, we feel it is safer to consider these families as independent until a stronger evidence for their grouping is available.

Bernard HENRISSAT*‡ and Amos BAIROCH†

*Centre de Recherches sur les Macromolécules Végétales§, C.N.R.S., BP 53, 38041 Grenoble Cédex, France, and †Medical Biochemistry Department, Centre Médical Universitaire, CH-1211 Geneva 4, Switzerland

‡ To whom correspondence should be addressed.
§ Affiliated with the Université Joseph Fourier, Grenoble, France.

1 Henrissat, B. (1991) Biochem. J. **280**, 309–316
2 Henrissat, B. and Bairoch, A. (1993) Biochem. J. **293**, 781–788
3 Appel R. D., Bairoch A. and Hochstrasser D. F. (1994) Trends Biochem. Sci. **19**, 258–260
4 Bairoch A. and Apweiler R. (1996) Nucleic Acids Res. **24**, 21–25
5 Bairoch A. (1996) Nucleic Acids Res. **24**, 221–222
6 Bairoch A., Bucher P. and Hofmann K. (1996) Nucleic Acids Res. **24**, 189–196
7 Sinnott, M. L. (1990) Chem. Rev. **90**, 1171–1202
8 Gebler, J. C., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Béguin, P., Wakarchuk, W. W., Kilburn, D. G., Miller, Jr., R. C., Warren, R. A. J. and Withers, S. G. (1992) J. Biol. Chem. **267**, 12559–12561
9 Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P. and Davies, G. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 7090–7094
10 Davies, G. and Henrissat, B. (1995) Structure **3**, 853–859
11 Jenkins, J., Lo Leggio, L., Harris, G. and Pickersgill, R. (1995) FEBS Lett. **362**, 281–285
12 Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T. and Jones, T. A. (1994) Science **265**, 524–528
13 Törrönen, A., Kubicek, C. P. and Henrissat, B. (1993) FEBS Lett. **321**, 135–139
14 Dagnall, B. H., Paulsen, I. T. and Saier, Jr., M. H. (1995) Biochem. J. **311**, 349–350
15 Romeu, A. and Henrissat, B. (1995) Biochem. J. **311**, 350–351
16 Crennel, S. J., Garman, E. F., Laver, W. G., Vimr, E. R. and Taylor, G. L. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 9852–9856
17 Holm, L. and Sander, C. (1994) FEBS Lett. **340**, 129–132

Received 11 March 1996